STRATEGI PENERAPAN UJI KEACAKAN BARISAN BIT MENGGUNAKAN UJI RUN

I Made Mustika Kerta Astawa

Lembaga Sandi Negara
kadek19_kaptainboy@yahoo.com

Abstrak

Uji run dapat digunakan untuk melihat apakah observasi (sampel) diambil secara random atau tidak. Barisan yang dilakukan pengujian menggunakan uji run dapat berupa uji biner maupun uji non biner. Pada penelitian ini dilakukan strategi penerapan untuk melakukan uji keacakan barisan bit menggunakan uji run. Misalnya satu kasus barisan biner dengan panjang n dari dua simbol, masing-masing sebanyak n1 dan n2, sehingga n=n1+n2. Jika r1 adalah banyaknya run untuk simbol 1 dan r2 adalah banyaknya run untuk simbol 2, maka total run adalah r=r1+r2. Untuk menurunkan statistik uji keacakan berdasar nilai R, maka diperlukan distribusi peluang dari peubah acak R tersebut pada kondisi hipotesis nol benar, yaitu barisan bersifat acak.

Kata kunci: *Uji Keacakan, Barisan Bit, Uji Run*

I. PENDAHULUAN

a) Latar Belakang

Perhatikan suatu barisan antrian dari orang-orang di suatu pintu masuk gedung bioskop. Ambil contoh ada 10 orang yang terdiri 5 perempuan (F) dan 5 laki-laki (M). Perhatikan dua barisan berikut :

1. M, F, M, F, M, F, M, F, M, F
2. M, M, M, M, M, F, F, F, F, F

Secara intuitif kita mengatakan bahwa kedua barisan tersebut tidaklah random. Yang pertama ada kemungkinan film yang diputar adalah film remaja, sehingga para penontonnya cenderung saling berpasangan. Sedangkan yang kedua kemungkinan adalah film anak-anak, sehingga penontonnya saling mengelompok sesuai jenis kelamin. Jika suatu barisan tertentu memberikan konsekuensi tertentu (seperti dua contoh di atas), maka kita katakan bahwa barisan tersebut tidaklah random.
Salah satu besaran yang bisa dipakai untuk mengindikasikan apakah suatu barisan random atau non random adalah 'run'.

Run adalah suatu segmen yang terdiri dari satu simbol yang tidak terputus yang diapit oleh dua simbol yang berbeda atau satu simbol dengan satu sisinya kosong.
Oleh karena itu, barisan (1) di atas terdiri 5 run untuk simbol M dan 5 run untuk simbol F, atau semuanya 10 run. Sedangkan untuk barisan (2) terdiri dari 1 run untuk simbol M dan 1 run untuk simbol F, sehingga totalnya ada 2 run.
Kalau kita perhatikan, barisan (1) terlalu banyak run, dan barisan (2) terlalu sedikit run. Hal ini yang menyebabkan secara intuitif kita memutuskan bahwa kedua barisan

tersebut tidaklah random. Dengan berbasis jumlah run tersebut, maka secara intuitif kita dapat mengatakan bahwa suatu barisan akan bersifat random jika banyaknya run tidak terlalu sedikit dan tidak terlalu banyak. Secara statistik, harus ada batas tertentu untuk mengambil salah satu dari 2 alternatif keputusan tersebut (random atau tidak random). Guna penentuan batas ini, maka harus diketahui distribusi dari jumlah run tersebut.

b) Tujuan

Tujuan uji run adalah untuk menentukan apakah jumlah run adalah sesuai yang diharapkan kalau barisan yang diberikan adalah random. Atau secara khusus untuk kasus biner, uji ini untuk menentukan apakah pergantian antara simbol 1 dan 0 adalah terlalu cepat (sehingga banyak run) atau terlalu lambat (sehingga sedikit run).

c) Metode Penelitian

Pada penelitian ini, penulis menggunakan metode penelitian kepustakaan berupa deskripsi penelitian yang dihasilkan atas kajian referensi pustaka dan eksperimen. Sama seperti bentuk penelitian lainnya, penelitian kepustakaan ini bertujuan untuk mengklarifikasi atau memperluas pemahaman dan pengetahuan. Tahapan proses kajian ini adalah sebagai berikut:

1. Pengumpulan data
   Melakukan pengumpulan referensi dari beberapa buku atau referensi lain mengenai uji keacakan barisan bit menggunakan uji run.
2. Implemetasi
   Melakukan proses pembuktian rumus-rumus statistik.
3. Analisis data
   Analisis hasil pengumpulan data dan kajian terhadap materi yang berkaitan dengan uji keacakan barisan bit menggunakan uji run.
4. Pengambilan Kesimpulan
   Pengambilan simpulan hasil penelitian.

II. PEMBAHASAN

a) Distribusi Exact Jumlah Run pada Barisan Biner

Ambil satu kasus barisan biner dengan panjang n dari dua simbol, masing-masing sebanyak n1 dan n2, sehingga n=n1+n2. Jika r1 adalah banyaknya run untuk simbol 1 dan r2 adalah banyaknya run untuk simbol 2, maka total run adalah r=r1+r2. Untuk menurunkan statistik uji keacakan berdasar nilai R, maka diperlukan distribusi peluang dari peubah acak R tersebut pada kondisi hipotesis nol benar, yaitu barisan bersifat acak.

Oleh karena R merupakan penjumlahan dari R1 dan R2, maka tahapannya dimulai dengan merumuskan distribusi bersama dari R1 dengan R2, kemudian distribusi marjinal masing-masing peubah dan baru kemudian distribusi jumlah dari peubah acak R1 dan R2 tersebut, yaitu R=R1+R2.

Berdasar definisi run, maka hanya ada 3 kemungkinan yang terjadi, yaitu :

$$r_1=r_2 \quad \text{atau} \quad r_1=r_2-1 \text{ atau} \quad r_1=r_2+1$$

Oleh karena itu, peluang bersama dari R1 dan R2 adalah :

ISBN: 978-602-70609-0-6
Tuban, 24 Mei 2014

$$f_{R1,R2}(r_1,r_2)=\frac{c\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_2-1}}{\binom{n_1+n_2}{n_1}}$$

Dengan r_1=1,2,3,...,n_1 dan c=2 untuk $r_1=r_2$ serta c=1 untuk $r_1=r_2\pm1$.

Pembagi pada fungsi peluang bersama tersebut merupakan banyaknya cara menyusun n unsur yang terdiri dari dua simbol, masing masing-masing sebanyak n1 dan n2. Sedangkan bagian atas merupakan perkalian antara banyaknya cara mengalokasikan n1 unsur ke dalam r1 kotak dengan ketentuan tidak boleh ada kotak yang kosong dengan banyaknya cara mengalokasikan n2 unsur ke dalam r2 kotak dengan ketentuan tidak boleh ada kotak yang kosong.

Oleh karena itu fungsi peluang masing-masing peubah acak R1 dan R2 adalah :

$$f_{R1}(r_1)=\sum_{\forall r_2} f_{R1,R2}(r_1,r_2)=\frac{2\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_1-1}+\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_1-2}+\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_1}}{\binom{n_1+n_2}{n_1}}$$

$$f_{R1}(r_1)=\frac{\binom{n_1-1}{r_1-1}\binom{n_2+1}{r_1}}{\binom{n_1+n_2}{n_1}}$$

Begitu juga untuk R2 :

$$f_{R2}(r_2)=\frac{\binom{n_2-1}{r_2-1}\binom{n_1+1}{r_2}}{\binom{n_1+n_2}{n_1}}$$

Oleh karena itu, fungsi peluang dari R=R1+R2, yaitu total banyaknya run dari sekuen sepanjang n=n1+n2, n1 untuk simbol 1 dan n2 untuk simbol 2 adalah :

$$f_R(r)=\begin{cases}\dfrac{2\binom{n_1-1}{r/2-1}\binom{n_2-1}{r/2-1}}{\binom{n_1+n_2}{n}} & r \quad genap\\[2ex] \dfrac{\binom{n_1-1}{(r-1)/2}\binom{n_2-1}{(r-3)/2}+\binom{n_1-1}{(r-3)/2}\binom{n_2-1}{(r-1)/2}}{\binom{n_1+n_2}{n}} & r \quad ganjil\end{cases}$$

Nilai r dalam hal ini adalah 2, 3, 4, ..., n.

Formula tersebut didasarkan fakta bahwa r genap hanya diperoleh untuk r1=r2=r/2. Oleh karena itu, jika r genap, maka fungsi peluang diperoleh dari persamaan (1) dengan mengganti r1 dan r2 dengan r/2.

Jika r ganjil, maka ada diperoleh dari 2 kemungkinan, yaitu r1=r2-1 atau mungkin juga r1=r2+1. Fungsi peluang R diperoleh dari (1) dengan menggantikan :

r1=(r-1)/2 dan r2=(r1+1)/2,

atau

r1=(r+1)/2 dan r2=(r-1)/2

Berdasar fungsi peluang R tersebut, maka kita bisa melakukan uji keacakan suatu barisan biner. Misalkan suatu sekuen simbol biner terdiri dari 5 simbo 1 dan 4 simbol 0, maka kita bisa melakukan perhitungan :

R=9 → $f_R(9)$=1/126=0.008

R=8 → $f_R(8)$=8/126=0.063

R=2 → $f_R(2)$=2/126=0.016

R=3 → $f_R(3)$=7/126=0.056

Untuk uji dua arah berikut :

Ho : sekuen adalah random

H1 : sekuen tidak random

Jika daerah penolakan Ho adalah R≤2 atau R≥9, maka taraf nyata uji adalah 0.016+0.008=0.024.

Keterbatasan dari distribusi exact tersebut dalam melakukan pengujian kerandoman adalah kesulitan menghitung untuk n yang sangat besar. Oleh karena itu untuk n yang snagat besar dilakukan dengan pendekatan distribusi normal. Untuk itu perlu diketahui rata-rata R, $E(R)$, dan variance dari R, $Var(R)$.

Nilai rata-rata dan variance ini bisa ditentukan dengan menggunakan moment ke k, yaitu:

$$E(R^k) = \sum_{\forall r} r^k f_R(r)$$

Nilai ini sama dengan :

$$\frac{\displaystyle\sum_{r \in genap} 2r^k \binom{n_1-1}{r/2-1}\binom{n_2-1}{r/2-1} + \sum_{r \in ganjil} r^k \left\{\binom{n_1-1}{(r-1)/2}\binom{n_2-1}{(r-3)/2} + \binom{n_1-1}{(r-3)/2}\binom{n_2-1}{(r-1)/2}\right\}}{\dbinom{n_1+n_2}{n_1}}$$

Nilai terkecil untuk r adalah 2. Jika n1=n2, maka nilai tertinggi dari run terjadi saat simbol-simbol muncul dalam sekuen saling bergantian, sehingga dalam hal ini nilai terbesar adalah r=2n_1. Jika n1<n2, maka nilai tertinggi r adalah 2n_1+1, yang berarti sekuen diawali dan diakhiri oleh simbol 2. Anggap $n_1 \le n_2$, maka nilai r adalah $2 \le r \le 2n_1+1$. Misal r=2i untuk r genap dan r=2i+1 untuk r ganjil, dengan $1 \le i \le n_1$. Dengan penjabaran seperti ini dan menggunakan formula moment di atas, maka diperoleh persamaan :

$$\binom{n_1+n_2}{n_1} E(R) = \sum_{i=1}^{n_1} 4i \binom{n_1-1}{i-1}\binom{n_2-1}{i-1} + \sum_{i=1}^{n_1} (2i+1)\binom{n_1-1}{i}\binom{n_2-1}{i-1} + \sum_{i=1}^{n_1} (2i+1)\binom{n_1-1}{i-1}\binom{n_2-1}{i}$$

Lemma 1 : $\displaystyle\sum_{r=0}^{c} \binom{m}{n}\binom{n}{r} = \binom{m+n}{m}$ dengan $c = \min\{m, n\}$

ISBN: 978-602-70609-0-6
Tuban, 24 Mei 2014

Bukti :

$$(1+x)^{m+n} = (1+x)^m(1+x)^n$$

Dengan menggunakan kaidah binom : $(a+b)^n = \sum_{i=0}^{n}\binom{n}{i}a^i b^{n-i}$, maka diperoleh :

$$\sum_{i=0}^{m+n}\binom{m+n}{i}x^i = \sum_{j=0}^{m}\binom{m}{j}x^j \sum_{k=0}^{n}\binom{n}{k}x^k$$

Dengan mengambil c=m, dan menyamakan ruas kiri dan kanan untuk koefisien x^m, maka diperoleh :

$$\binom{m+n}{m} = \sum_{r=0}^{m}\binom{m}{m-r}\binom{n}{r}$$

Lemma 2 : $\sum_{r=0}^{c}\binom{m}{r}\binom{n}{r+1} = \binom{m-n}{m-1}$ dengan $c = \min\{m, n-1\}$

Dengan kedua lemma ini maka nilai harapan R dapat diperoleh. Sedangkan untuk menghitung variance R dilakukan dengan mengevaluasi $E\left[R(R-1)\right]$.

Namun demikian ada cara mudah untuk menghitung harapan R dan variance R, yaitu dengan menggunakan peubah indikator.

Misalkan :

$$R=1+I_2+I_3+...+I_n$$

Dengan I_k didefinisikan sebagai :

$$I_k = \begin{cases} 1 & \varepsilon_k \neq \varepsilon_{k-1} \\ 0 & selainnya \end{cases}$$

Maka I_k adalah peubah acak Bernoulli dengan parameter :

$$p = \frac{2n_1n_2}{n(n-1)}$$

Oleh karena itu :

$$E(I_k) = E(I_k^2) = \frac{2n_1n_2}{n(n-1)}$$

Oleh karena R adalah kombinasi linear dari I_k, maka :

$$E(R) = 1 + \sum_{k=2}^{n}E(I_k) = 1 + \frac{2n_1n_2}{n}$$

$$\mathrm{var}(R) = \mathrm{var}\left(\sum_{k=2}^{n}I_k\right) = (n-1)\mathrm{var}(I_k) + \sum_{2\le j\ne k\le n}\sum \mathrm{cov}(I_jI_k)$$

$$= (n-1)E(I_k^2) + \sum_{2\le j\ne k\le n}\sum E(I_jI_k) - (n-1)^2[E(I_k)]^2$$

Dalam hal ini ada (n-1)(n-2) komponen $E(I_jI_k)$ untuk $j\neq k$, dengan indeks tersebut dapat dikelompokkan menjadi :

1. Ada 2(n-2) pasangan (j,k) dengan j=k-1 dan j=k+1

$$E(I_jI_k) = \frac{n_1n_2(n_1-1) + n_2n_1(n_2-1)}{n(n-1)(n-2)} = \frac{n_1n_2}{n(n-1)}$$

2. Pasangan sisanya sebanyak : (n-1)(n-2)-2(n-2)=(n-2)(n-3) berlaku

$$E(I_j I_k) = \frac{4n_1 n_2 (n_1 - 1)(n_2 - 1)}{n(n-1)(n-2)(n-3)}$$

Dengan mengsubstitusikan keduanya ke dalam persamaan variance di atas maka diperoleh :

$$\begin{aligned} \operatorname{var}(R) &= \frac{2n_1 n_2}{n} + \frac{2(n-2)n_1 n_2}{n(n-1)} + \frac{4n_1 n_2 (n_1 - 1)(n_2 - 1)}{n(n-1)} - \frac{4n_1^2 n_2^2}{n^2} \\ &= \frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)} \end{aligned}$$

b) Distribusi Limit Jumlah Run pada Barisan Biner

Pada bagian sebelumnya telah diperoleh distribusi exact dari jumlah run, R. Dengan distribusi ini kita bisa menggunakannya untuk menguji kerandoman suatu barisan. Namun demikian, untuk sekuen yang sangat panjang, maka ada kendala dalam melakukan penghitungan. Oleh karena itu perlu didekati dengan distribusi limit.

Untuk n yang sangat besar, $n \to \infty$, maka berlaku :

$$\frac{n_1}{n} \to \lambda \text{ dan } \frac{n_2}{n} \to 1 - \lambda \text{ dengan } 0<\lambda<1$$

Dengan demikian, rata-rata dan variance menjadi :

$$\lim_{n\to\infty} E\left(\frac{R}{n}\right) = 2\lambda(1-\lambda) \quad \text{dan} \quad \lim_{n\to\infty} \operatorname{var}\left(\frac{R}{\sqrt{n}}\right) = 4\lambda^2(1-\lambda)^2$$

Dari kedua besaran tersebut, maka diperoleh peubah acak normal baku sebagai berikut :

$$Z = \frac{R - 2n\lambda(1-\lambda)}{2\sqrt{n}\lambda(1-\lambda)}$$

Uji 2 arah untuk keacakan dengan jumlah run menggunakan pendekatan distribusi Normal akan menolak Ho dan menyimpulkan bahwa sekuen adalah tidak random kalau dipenuhi :

$$\left| \frac{R - 2n\lambda(1-\lambda)}{2\sqrt{n}\lambda(1-\lambda)} \right| \geq z_{\alpha/2}$$

c) Strategi Penerapan Untuk Kasus Non Biner

Misalkan dalam barisan yang akan diuji terdiri dari K simbol, yaitu 1, 2, 3, ..., K, maka mengikuti perumusan sebelumnya, uji run test untuk barisan non biner dapat diturunkan dengan alur logika sebagai berikut :

Pada suatu barisan yang akan diuji dengan panjang n, dengan masing-masing simbol terdiri sebanyak n_1, n_2, ..., n_k, dengan

$$n_1 + n_2 + n_3 + \ldots + n_K = n$$

maka jumlah run dapat dirumuskan sebagai :

$$R = 1 + I_2 + I_3 + \ldots + I_n$$

ISBN: 978-602-70609-0-6
Tuban, 24 Mei 2014

Dalam hal ini I_k didefinisikan sebagai :

$$I_k = \begin{cases} 1 & \varepsilon_k \neq \varepsilon_{k-1} \\ 0 & selainnya \end{cases}$$

Maka I_k adalah peubah acak Bernoulli dengan parameter :

$$p = \frac{n_1(n-n_1)+n_2(n-n_2)+n_3(n-n_3)+\ldots+n_K(n-n_K)}{n(n-1)} = \sum_{i=1}^{K}\frac{n_i(n-n_i)}{n(n-1)}$$

Oleh karena itu :

$$E(I_k) = E(I_k^2) = \sum_{i=1}^{K}\frac{n_i(n-n_i)}{n(n-1)} = \frac{1}{n(n-1)}\sum_{i=1}^{K} n_i(n-n_i)$$

Oleh karena R adalah kombinasi linear dari I_k, maka :

$$E(R) = 1+\sum_{k=2}^{n}E(I_k) = 1+\frac{1}{n}\sum_{i=1}^{K}n_i(n-n_i) = 1+\frac{n\sum_{i=1}^{K}n_i - \sum_{i=1}^{K}n_i n_i}{n} = 1+\frac{\sum_{i=1}^{K}\sum_{j\neq i, j=1}^{K}n_i n_j}{n}$$

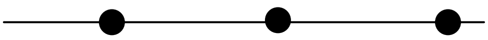
$$\text{var}(R) = \text{var}\left(\sum_{k=2}^{n} I_k\right) = (n-1)\text{var}(I_k) + \sum_{2\le j\neq k\le n}\sum \text{cov}(I_j I_k)$$

$$= (n-1)E(I_k^2) + \sum_{2\le j\neq k\le n}\sum E(I_j I_k) - [(n-1)+(n-1)(n-2)][E(I_k)]^2$$

$$= (n-1)E(I_k^2) + \sum_{2\le j\neq k\le n}\sum E(I_j I_k) - (n-1)^2[E(I_k)]^2$$

Dalam hal ini ada (n-1)(n-2) komponen $E(I_j I_k)$ untuk $j \neq k$, dengan indeks tersebut dapat dikelompokkan menjadi :

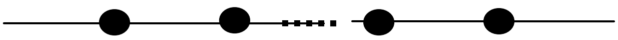
1. Ada 2(n-2) pasangan (j,k) dengan bentuk j=k-1 atau bisa juga j=k+1



	j	k
n_i	n_j dg $j\neq i$	(n_i-1+n-n_j-n_i) =(n-n_j-1)
n	n-1	n-2

$$E(I_j I_k) = \frac{\sum_{i=1}^{K}\sum_{j\neq i, j=1}^{K} n_i n_j (n-n_j-1)}{n(n-1)(n-2)}$$

2. Pasangan sisanya sebanyak : (n-1)(n-2)-2(n-2)=(n-2)(n-3) berlaku



	j		k
n_i	n-n_i	n_i-1	n-n_i-1
n_i	n-n_i	n-n_i-1	n_i-1
n	n-1	n-2	n-3

$$E(I_j I_k) = \frac{\sum_{i=1}^{K} 2n_i(n-n_i)(n_i-1)(n-n_i-1)}{n(n-1)(n-2)(n-3)}$$

Dengan mengsubstitusikan keduanya ke dalam persamaan variance di atas maka diperoleh :

$$\operatorname{var}(R)=\frac{2n_1n_2}{n}+\frac{2(n-2)n_1n_2}{n(n-1)}+\frac{4n_1n_2(n_1-1)(n_2-1)}{n(n-1)}-\frac{4n_1^2n_2^2}{n^2}$$

$$=\frac{2n_1n_2\left(2n_1n_2-n_1-n_2\right)}{(n_1+n_2)^2(n_1+n_2-1)}$$

$$\operatorname{var}(R)=(n-1)E\left(I_k^2\right)+\sum_{2\le j\ne k\le n}\sum E\left(I_jI_k\right)-(n-1)^2\left[E\left(I_k\right)\right]^2$$

$$\operatorname{var}(R)=(n-1)\frac{1}{n(n-1)}\sum_{i=1}^{K}n_i\left(n-n_i\right)+\sum_{2\le j\ne k\le n}\sum E\left(I_jI_k\right)-(n-1)^2\left[\frac{1}{n(n-1)}\sum_{i=1}^{K}n_i\left(n-n_i\right)\right]^2$$

a. $$\sum_{2\le j\ne k\le n}\sum E\left(I_jI_k\right)=2(n-2)\frac{\sum_{i=1}^{K}\sum_{j\ne i,j=1}^{K}n_in_j\left(n-n_j-1\right)}{n(n-1)(n-2)}+(n-2)(n-3)\frac{\sum_{i=1}^{K}2n_i\left(n-n_i\right)\left(n_i-1\right)\left(n-n_i-1\right)}{n(n-1)(n-2)(n-3)}$$

$$=\frac{2\sum_{i=1}^{K}\sum_{j\ne i,j=1}^{K}n_in_j\left(n-n_j-1\right)}{n(n-1)}+\frac{\sum_{i=1}^{K}2n_i\left(n-n_i\right)\left(n_i-1\right)\left(n-n_i-1\right)}{n(n-1)}$$

b. $$(n-1)\frac{1}{n(n-1)}\sum_{i=1}^{K}n_i\left(n-n_i\right)=\frac{1}{n}\sum_{i=1}^{K}n_i\left(n-n_i\right)=\sum_{i=1}^{K}\frac{n_i\left(n-n_i\right)}{n}$$

c. $$-(n-1)^2\left[\frac{1}{n(n-1)}\sum_{i=1}^{K}n_i\left(n-n_i\right)\right]^2=-\left[\sum_{i=1}^{K}\frac{n_i\left(n-n_i\right)}{n}\right]^2$$

maka Var(R) adalah :

$$\operatorname{var}(R)=\sum_{i=1}^{K}\frac{n_i\left(n-n_i\right)}{n}+\frac{2\sum_{i=1}^{K}\sum_{j\ne i,j=1}^{K}n_in_j\left(n-n_j-1\right)}{n(n-1)}+\frac{\sum_{i=1}^{K}2n_i\left(n-n_i\right)\left(n_i-1\right)\left(n-n_i-1\right)}{n(n-1)}-\left[\sum_{i=1}^{K}\frac{n_i\left(n-n_i\right)}{n}\right]^2$$

Sebagai contoh untuk kasus Biner :

$$\operatorname{var}(R)=\sum_{i=1}^{2}\frac{n_i\left(n-n_i\right)}{n}+\frac{2\sum_{i=1}^{2}\sum_{j\ne i,j=1}^{2}n_in_j\left(n-n_j-1\right)}{n(n-1)}+\frac{\sum_{i=1}^{2}2n_i\left(n-n_i\right)\left(n_i-1\right)\left(n-n_i-1\right)}{n(n-1)}-\left[\sum_{i=1}^{2}\frac{n_i\left(n-n_i\right)}{n}\right]^2$$

$$\operatorname{var}(R)=\frac{2n_1n_2}{n}+\frac{2(n-2)n_1n_2}{n(n-1)}+\frac{4n_1n_2(n_1-1)(n_2-1)}{n(n-1)}-\frac{4n_1^2n_2^2}{n^2}$$

Seperti pada kasus biner, maka pada non biner untuk barisan yang panjang juga didekati dengan distribusi limit.

Untuk n yang sangat besar, $n \to \infty$, maka berlaku :

$$\frac{n_k}{n} \to \lambda_k \text{ untuk k=1, 2, 3, ..., K dengan } 0<\lambda<1 \text{ dan } \sum_{i=1}^{K} \lambda_i = 1$$

Dengan demikian, rata-rata dan variance menjadi :

$$\lim_{n\to\infty} E\left(\frac{R}{n}\right) = \lim_{n\to\infty} \frac{1}{n}\left(1 + \frac{\sum_{i=1}^{K}\sum_{j\neq i, j=1}^{K} n_i n_j}{n}\right) = \lim_{n\to\infty} \sum_{i=1}^{K} \frac{1}{n^2} - \lim_{n\to\infty} \sum_{i=1}^{K}\sum_{j\neq i, j=1}^{K} \frac{n_i}{n}\frac{n_j}{n} = 2\sum_{i=1}^{K-1}\sum_{j=i+1}^{K} \frac{n_i n_j}{n^2} = 2\sum_{i=1}^{K-1}\sum_{j=i+1}^{K} \lambda_i \lambda_j$$

dan

$$\lim_{n\to\infty} \text{var}\left(\frac{R}{\sqrt{n}}\right) = \ldots$$

Dari kedua besaran tersebut, maka diperoleh peubah acak normal baku sebagai berikut :

$$Z = \frac{R - \ldots.}{\ldots\ldots.}$$

Uji 2 arah untuk keacakan dengan jumlah run menggunakan pendekatan distribusi Normal akan menolak Ho dan menyimpulkan bahwa sekuen adalah tidak random kalau dipenuhi :

$$\left|\frac{R - \ldots.}{\ldots\ldots.}\right| \geq z_{\alpha/2}$$

III. KESIMPULAN

Uji run secara khusus untuk kasus biner, uji ini untuk menentukan apakah pergantian antara simbol 1 dan 0 adalah terlalu cepat (sehingga banyak run) atau terlalu lambat (sehingga sedikit run). Uji run juga dapat diterapkan untuk menguji keacakan barisan bit non biner. Strategi yang diterapkan untuk penggunaan barisan bit antara barisan bit biner dan non biner hampir sama.

IV. DAFTAR PUSTAKA

Ghahramani, S. (2005). Fundamentals of Probability with Stochastic Process. Third Edition. Prentice Hall, New Jersey

Chung, K. L. (2001). A Course in Probability Theory. Third Edition. Academic Press,San Diego.

Andrew Rukhin, Juan Soto, James Nechvatal, Miles Smid, Elaine Barker, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks,Alan Heckert, James Dray, San Vo (2010). A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Special Publication 800-22 National Institute of Standard and Technology.